

**REMARKS**

Claim 5 has been canceled. Claims 4 and 6-9 have been amended. Claims 1-4 and 6-10 remain in the application.

In the previous office action, claims 1-3 were allowed, and claims 8-10 were identified as being drawn to allowable subject matter.

In order to place the application in *prima facie* condition for allowance, claim 4 has been amended to include some (but not all) of the elements of claims 5, 8, and 9.

Claims 4-7 were rejected as being anticipated by Chen (EPO 0674315). Amended claim 4, as amended, is distinguished from the Chen reference. In particular, claim 4 now requires the step of using the new vocabulary to generate a line alignment of the audio speech signal with a corresponding word of the base language vocabulary. This step is similar to that presented in claims 8 and 9, which were identified to be allowable, and is not found in the Chen reference.

The present invention matches audio data in one language with a phone in other language. Such mapping is not mentioned in Chen. The method employed does not require the novel language to be recognized or translated into phonemes. In contrast, Chen explicitly says that the speech signal analysis is performed by a speech recognition system (col 4, line 58). Chen does not mention whether or how a conventional speech recognition can handle speech in a language that is other than the one for which the speech recognition system exists. The present invention specifically provides for "implementing audio driven facial animation system in a first language, referred to as the novel language using a speech recognition system of a second language, referred to as the base language." This frees the system from the requirement of a speech recognition system of the first or novel language, unlike Chen, which requires functional speech recognition for both languages.

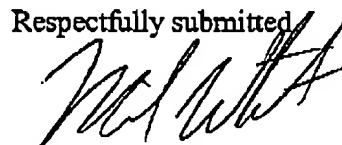
As explained on page 12 of the application specification, a particular advantage of the invention is that it eliminates the need to build a speech recognition engine for the same language in which the visual speech is to be synthesized. That is, given speech recognition for one language, using this

5

invention, one can obtain a synthesized video in another language. Further, the viseme images can be of only the language of which the alignment generation system is built, thereby eliminating the need for generating new viseme images for each language.

In view of the above, claims 1-4 and 6-10 should now be in *prima facie* condition for allowance. Reconsideration at an early date is requested.

Respectfully submitted



Michael E Whitham  
Reg. No. 32,635

Whitham, Curtis & Christofferson, P.C.  
11491 Sunset Hills Road, Suite 340  
Reston, VA 20190

Tel. (703) 787-9400  
Fax. (703) 787-7557

I certify that this five (5) page amendment was filed in the U.S.P.T.O. on October 30, 2003, by transmission to the facsimile having telephone number 703-872-9314.



Michael E. Whitham